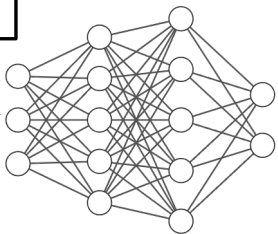


Reward r

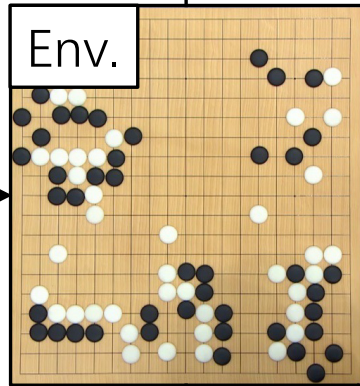
Agent



Policy $\pi_{DNN}(s, a)$

Action a

Env.



State s